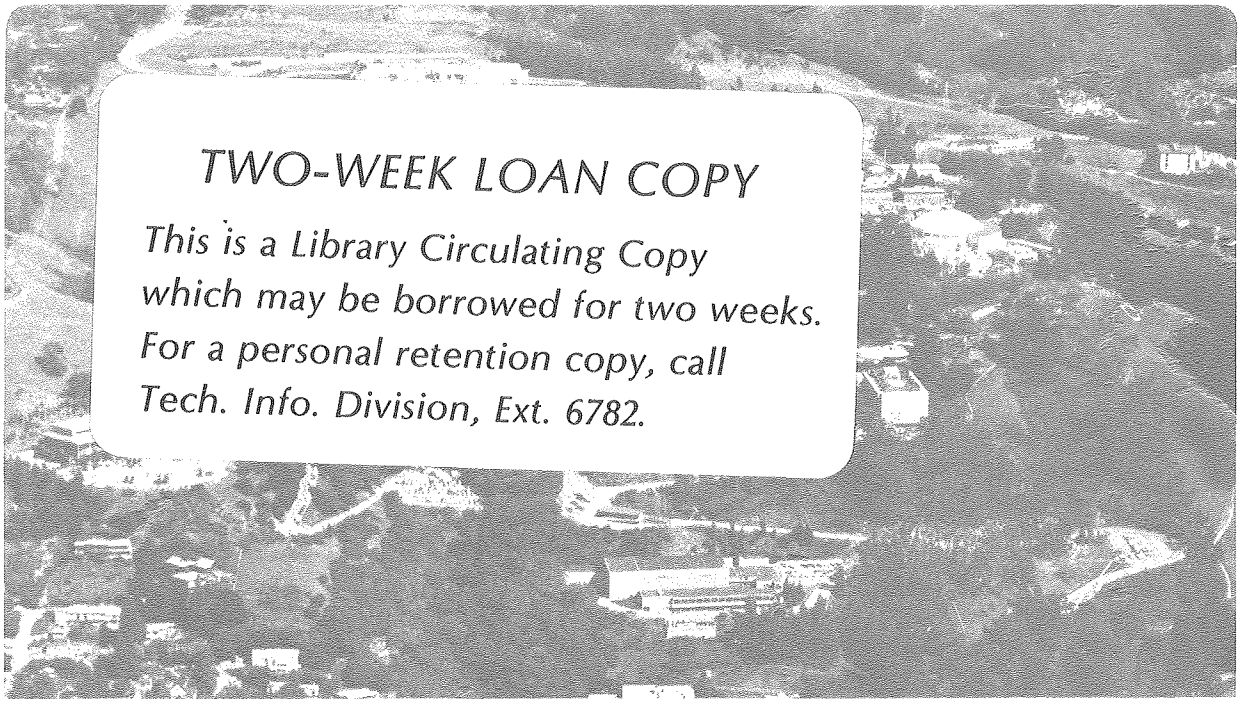
Lawrence Berkeley Laboratory
UNIVERSITY OF CALIFORNIA

Engineering & Technical
Services Division

AN INTERACTIVE SPIRES PLOT PROTOCOL

Barbara Cerny, Gerald Litton, and
J. Dennis Lawrence

June 1980



Prepared for the U.S. Department of Energy under Contract W-7405-ENG-48

DISCLAIMER

LBL-11068

AN INTERACTIVE SPIRES PLOT PROTOCOL

by

Barbara Cerny, Gerald Litton and J. Dennis Lawrence

June , 1980

Information and Data Analysis Department
Lawrence Berkeley Laboratory
University of California
Berkeley, California 94720

This work was supported by the U.S. Department of Energy under Contract Number W-7405-ENG-48.

Introduction

This report describes PLOT, a routine to plot a two-dimensional array of points from elements in any SPIRES database. Its use assumes access to a SPIRES database and knowledge of SPIRES querying and searching procedures. PLOT is not meant to be a replacement for more sophisticated on-line or off-line graphics packages compatible with SPIRES, but rather it gives immediate interactive plots that can be displayed at the terminal, saved, or disposed to a line printer. One can select records from a database, plot them, add other points, or change the scale, if desired, and replot in order to get a feeling for the form and distribution of the data. PLOT can be used on either a hard-copy or CRT terminal, and it allows user adjustment of various parameters to generate a plot compatible with both the data and the terminal display capabilities. The search commands that generate the stack of data to be plotted are identical to those used in any other SPIRES search. Both indexed and non-indexed records can be accessed and plotted, and elements can be obtained from any depth within structures.

PLOT is available to users of the SPIRES system at Stanford, as described in Appendix I.

A Few Words About SPIRES

SPIRES (Stanford Public Information Retrieval System) is a generalized database management system developed by Stanford University that went into operation in 1972 on the IBM 360/67 at the Stanford Center for Information Processing. Because of the flexibility of the design, the system has been installed at more than 25 institutions to date. The applications range from the enormous RLIN (Research Libraries Information Network) project at Stanford, to administrative uses such as the management of student records, to special purpose databases for the control of bibliographic and numeric data. Currently, at LBL, these latter areas are represented by databases in such diverse fields as Ventilation and Air Quality, Geothermal Resources and Mutagenesis.

The systems aspects of transforming data from user to machine representation and the structural design of SPIRES are described in (1). From a user perspective, the first step in entering data into the SPIRES system is to define a file using a special file-definition language (2). This specifies how the data are to be organized (logically, not structurally), how they are to be indexed and searched, and who can use the file. The data are organized into records which are made up of a variety of elements that can be fixed or varying in length, required or optional, singly or multiply occurring. As records are added, SPIRES builds indexes so that these elements can be searched. (The option also exists to do sequential, i.e., unindexed, searching).

The user who has defined, or has access to, a file can retrieve records by a query or search procedure and display them with a standard or customized format, or "report generator". Special purpose report generators such as mailing labels, catalogues, letters, lists of names, etc. are frequently used, and the PLOT routine adds another dimension to this phase of a user's interaction with databases.

The PLOT Routine

The plot routine was written using two SPIRES system-supported languages, Protocols language (3) and Format-definition language (4). (The File-definition language mentioned above is the third such language). Protocols language, being an extension of the command language, is a set of SPIRES, WYLBUR, MILTEN and ORVYL commands. It is particularly useful for generating prompts whose answers assume no knowledge of SPIRES commands on the part of the user. The prompt responses in PLOT control the flow of logic through the various plot options and issue calls to routines written in Format-definition language. These routines obtain the elements in the records containing the points, scale the axes and then plot the points. (See Appendix II for more technical detail).

The basic flow through PLOT is shown in Appendix III. The routine is self-contained in that instructions for use are given in response to prompts throughout execution. The user is prompted for the name of a specific database and then the elements of that database which are to be plotted. If the elements do not occur at the record level, then execution efficiency can be enhanced by specifying the pathname through the structure. For example, in LBL's Geothermal database one might wish to plot pressure vs. temperature. These occur in structures with the following pathnames:

Wellhead-data @ Flow-data @ Pressure
Wellhead-data @ Flow-data @ Temperature

If, in response to the prompt for element names, only "pressure" and "temperature" were given, PLOT would exhaustively check all paths through the subfile and return information on the pathnames leading to those particular elements. If "pressure" and "temperature" were multiply occuring elements names, e.g.

1. Wellhead-Data @ Flow-Data @ Pressure
2. Down-Hole-Data @ Pressure-Data @ Pressure

then both pathnames would be returned and the user prompted for the one he wished to plot.

Figure 1 shows the plot options available to the user

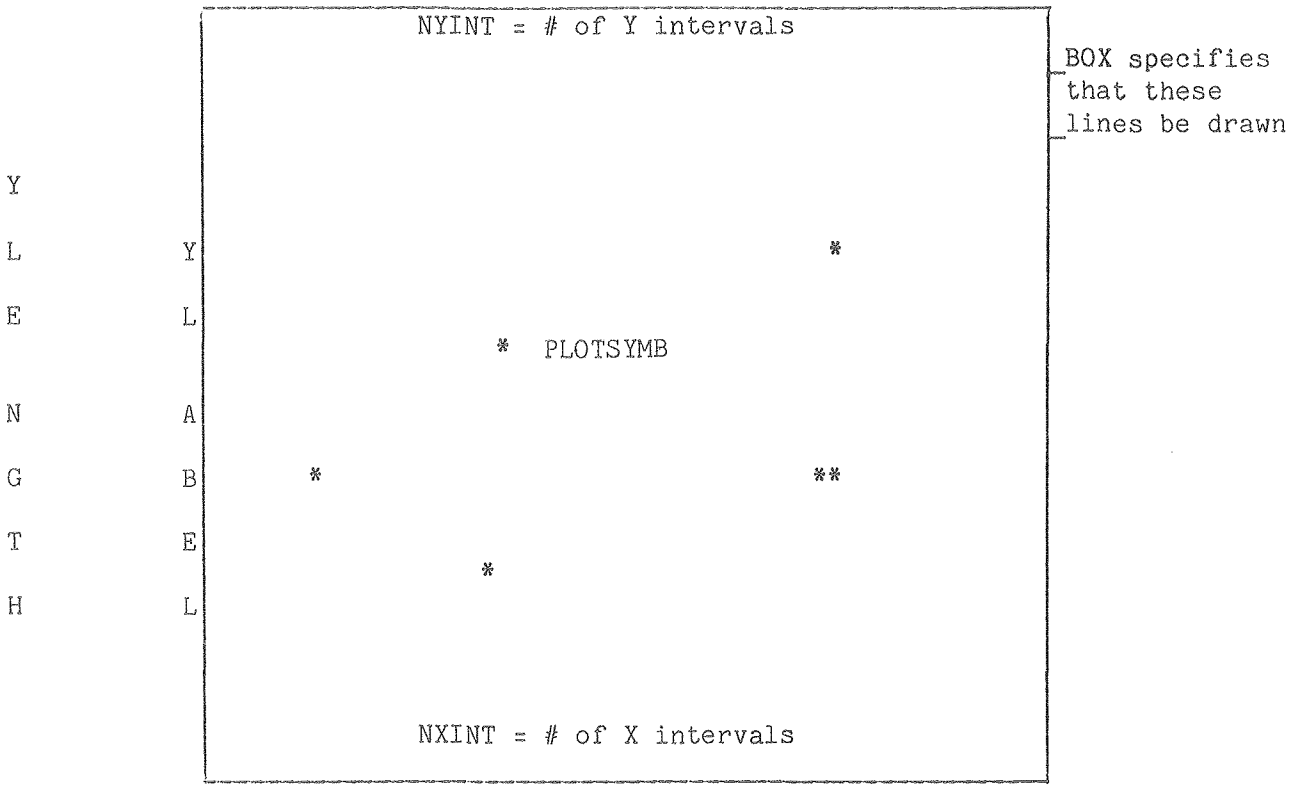


FIGURE 1

The default values and a brief description of these user controlled variables follows:

XLENGTH = 80
XLENGTH is the length of the X-axis (maximum is 130).

YLENGTH = 23
YLENGTH is the length of the Y-axis (maximum is 125).
N.B. XLENGTH and YLENGTH can only be approximately specified, since the plot routine may recalculate them to optimize intervals of the abscissa and ordinate.

NXINT = 10
NXINT is the number of intervals into which the X-axis will be divided.

NYINT = 5

NYINT is the number of intervals into which the Y-axis will be divided. N.B. Depending on the maximum and minimum values of the data elements, the number of intervals internally calculated approximate, but do not necessarily equal, NXINT and NYINT.

XSP = 8

XSP is the number of spaces into which the Y-axis label and ordinate scale are placed. That is, it is the number of spaces along the X-axis before the ordinate is drawn (minimum is 3).

PLOTSYMB = *

PLOTSYMB is the symbol that will be used to plot the data points.

XLABEL = The label assigned to the X-axis. Initially, this is set to the name of the X-element.

YLABEL = The label assigned to the Y-axis. Initially, this is set to the name of the Y-element.

N.B. If the X and Y elements had been specified as "STRUC1@STRUC2@XL" and "STRUC1@STRUC2@YL"
Then XLABEL = XL
and YLABEL = YL
These label names may be changed by the user.

BOX = YES

BOX specifies that lines parallel to the ordinate and abscissa will be drawn at the maxima of the X- and Y- scales. This is particularly useful when the data points are not confined to Quadrant I.
BOX = NO can be set.

Since a user might have access to either a CRT or a hard-copy terminal, there is an option for putting the plot in the active file or displaying it locally at the terminal. For quick trial plots on a CRT to display the data or to determine a good scale, a local plot might suffice. However, if the user wished to save the plot, it could be placed in the active file, then saved and disposed to a line printer.

Error messages which do not allow a plot to be generated occur in several cases.

1. If the data are non-numeric.
2. If there are unequal number of X and Y data points.
3. If either of the following occurs: (XMAX - XMIN = 0) or (YMAX - YMIN = 0). This situation makes it impossible for PLOT to determine a scale for the X and Y variables.
4. If no records are found.
5. If an invalid SPIRES search command is issued.

Appendix IV gives an example of a plot from a Geothermal database.

REFERENCES

1. Schoeder, J.F., Kiefer, W.C., Guertin, R.L., Berman, W.J., Stanford's Generalized Database System. Proceedings of the Conference on Very Large Database; Sept. 1975, pp. 120-43.

2. SPIRES File definition, SCIP Campus Computing Facility. Stanford University. Updated every few months.

3. SPIRES Protocols, SCIP Campus Computing Facility. Stanford University. Updated every few months.

4. SPIRES formats, SCIP Campus Computing Facility. Stanford University. Updated every few months.

APPENDIX I

To access PLOT on the Stanford SPIRES system.

```
SET  COMPXEQ PLOT
     ..CPLOT
```

APPENDIX II

A few technical details

PLOT consists of one protocol and two formats

Protocol

The protocol controls the flow of logic by using the responses to user prompts. These responses 1) Operate locally within the protocol; 2) "Drop through" to SPIRES to execute commands in SPIRES control language (For example,

```
SELECT [subfile]

SHOW ELEMENT NAMES
```

are generated by prompt responses while

```
FILE [index]
FOR TREE
etc.
```

or other SPIRES search commands are entered directly by the user); 3) Call the formats described below.

Extensive use is made of the SPIRES functions "ELEMTEST" and "ELNOTEST," which return the type of the element that a user has indicated that he wishes to plot. It checks for legal element names and then whether or not the element is numeric, i.e., real or integer. Most importantly, however, by doing a recursive "tree walk" through the structure of the database and using "ELNOTEST" at every element, it is possible to obtain the complete pathname for any element at any depth in the structure. This information is then used by a format to obtain the points.

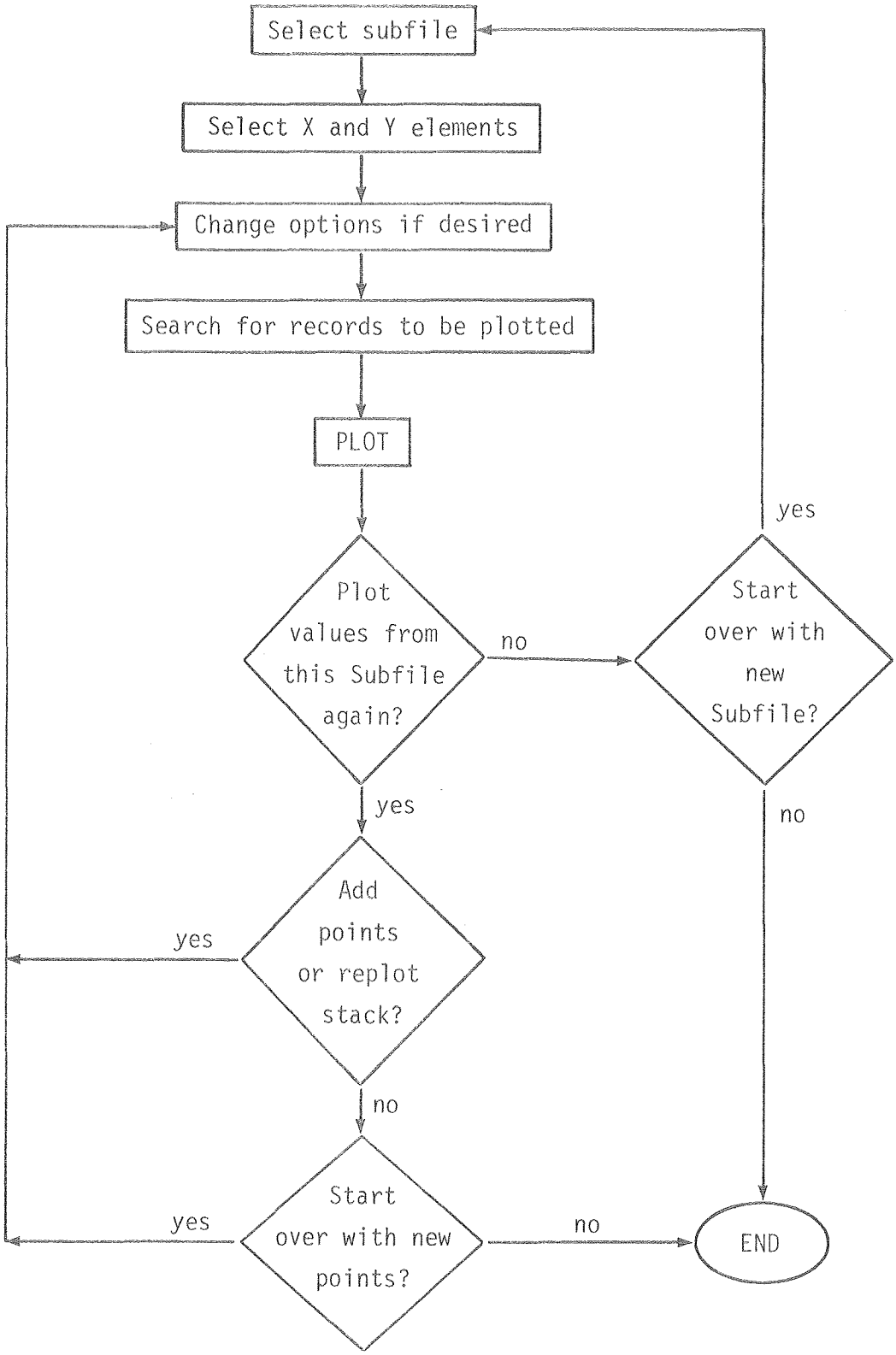
Formats

1. The first format obtains the points from elements in the database. If the element is at the record level, the process is trivial. If it is within a structure, one recursive frame controls the process of obtaining an element at any depth.

2. The second format calculates the axis scales and places the points. An algorithm is used that utilizes the user-entered parameter values or default values, as explained above, to calculate the axes scales. If the values of X and Y lie between 10 and 100, the axes are not scaled; otherwise, they are expressed in exponent form. The data point positions are calculated as a fraction of the distance between the MIN and MAX values of X and Y as calculated by the axis-scaling subroutine. Since this is a format, the actual entry of the points is accomplished simply by a LABEL - GROUP within a dimensioned frame.

```
LABEL;
VALUE = # PLOTSYMB;
START = # YPOSITION, #XPOSITION;
PUTDATA;
```

APPENDIX III

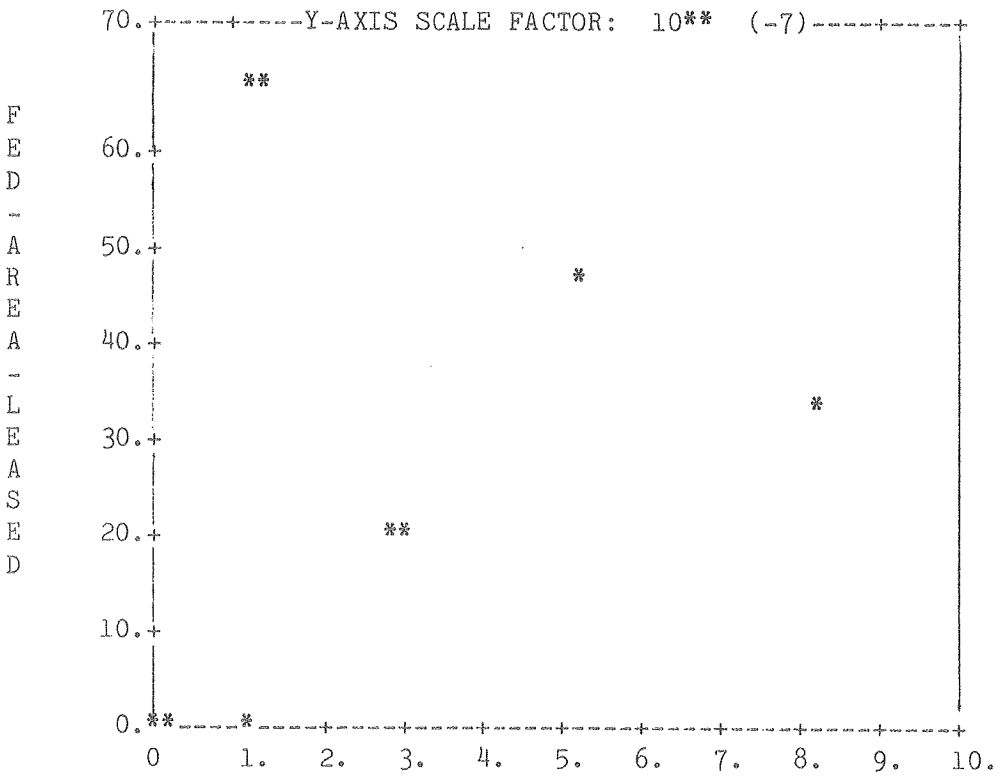


APPENDIX IV

Sample Plot

```
*Do you wish an explanation of this routine? (Y/N)
:N
*
*Do you wish your plots to appear at your terminal or be placed in your
*active file?  (CRT/ACTIVE)
*
:CRT
*
*From which subfile do you wish to plot?
:SITE SUMMARY
*
*Do you wish to see the elements in your subfile? (Y/N)
:N
*
*X-axis element?
:NBR-FED-LEASES
*
*Y-Axis element?
:FED-AREA-LEASES
*"FED-AREA-LEASES" is not a legal element name; please try again
:FED-AREA-LEASED
*
*The Y-element you requested exists in the following structure:
*
*                    LEASING-PROPO@FED-AREA-LEASED
*
*
*
*The X-element must come from the same structure.
*Axis label names = Axis element names? (Y/N/?)
:Y
*
*
*View current options? (Y/N)
:Y
*                          ---CURRENT PLOT OPTIONS---
*
*
*                               XLENGTH = 80
*                               YLENGTH = 23
*                               NXINT = 10
*                               NYINT = 5
*                               XLABEL = NBR-FED-LEASES
*                               XLABEL = FED-AREA-LEASED
*                               XSP = 8
*                               PLOTSYMB = *
*                               BOX = YES
```

```
*Enter new options?  (Y/N)? - use ? for instructions)
:Y
:LET YLENGTH = 25
:LET XLENGTH = 60
: ***
*View new options?  (Y/N)
:N
*
*Please enter whatever SEARCH-COMMANDS are necessary to get all records from
*  which data is to be plotted.  When search is finished,
-TYPE 'CONTINUE XEQ' TO RESUME
X-  FIND STATE CA
-RESULT: 21 SITE(s)
X-  OR STATE NY OR NM OR ID
-RESULT: 38 SITE(s)
X-  CONTINUE XEQ
-STACK: 38 SITE(s)
-END OF GLOBAL FOR
*
*
*
```



```
*Do you wish to plot values from this subfile again (Y/N/ACTIVE/?)
:N
*
*End or select new subfile? (END/SELECT)
:E
*DONE:
```